Supplementary: Boundary Multiple Measurement Vectors for Multi-Coset Sampler

Dong Xiao, Jian Wang and Yun Lin

This supplementary material is dedicated to the proofs for Theorems 1–3 in our main paper.

Before proceeding to the proofs, we review some useful notations. For a complex matrix $\mathbf{X} \in \mathbb{C}^{n\times L}$ and a set $S \subseteq \{1, \cdots, n\}$, $\mathbf{X}_S$ (or $\mathbf{X}^S$) denotes the submatrix of $\mathbf{X}$ with columns (or rows) indexed by S; $\mathbf{X}_{i,j}$, $\mathbf{X}_{i,:}$ and $\mathbf{X}_{:,i}$ are the (i, j)th entry, ith row and ith column of $\mathbf{X}$, respectively; $\mathbf{X}^{\dagger}$, $\mathbf{X}^{H}$ and $\mathbf{X}^{\top}$ mean the Moore-Penrose pseudo-inverse, conjugate transpose and transpose of $\mathbf{X}$, respectively; $\text{supp}(\mathbf{X})$ is the non-zero row indices (i.e., joint sparsity) of $\mathbf{X}$; $\|\mathbf{X}\|_F$ and $\|\mathbf{X}\|_2$ signify the Frobenius and Euclidean norm of $\mathbf{X}$, respectively. Moreover, S^c is the complement of set S; $\mathbf{I}_L$ is an $L \times L$ identity matrix.

I. Proof of Theorem 1

Theorem 1. *The actual sampling rate of (4) is* $\min\left(pf_s, f_{nyq}\right)$, *which attains the theoretical lower bound of sampling rate in MCS when* $|supp(\mathbf{X})| \leq \frac{N_{sig}B}{f_s}$.

Proof. Assume that a multi-coset sampler has p channels. In the ith channel, The sampling sequence is given by

$$x_{c_i}[n] = x(LTn + \tau_i), \quad n = 0, 1, \cdots . \tag{S.1}$$

The sampling rate of each channel is determined by the sampled signal sequence. Because the sampling time interval is LT, the sampling rate of each channel is 1 in L of the Nyquist sampling rate (i.e. $\frac{f_{\text{nyq}}}{L}$). The average sampling rate of p channels is p times that of one channel (i.e. $\frac{pf_{\text{nyq}}}{L}$). Noting that when the sampling rate of p channels is greater than the Nyquist rate, the advantage of sub-Nyquist sampling structure no longer exists. Thus, we only need to sample at Nyquist sampling rate f_{nyq}, the actual sampling rate can be represented as

$$\min\left(\frac{p}{LT}, f_{\text{nyq}}\right). \tag{S.2}$$

The theoretical lower bound of the sampling rate is given in [17], which is directly determined by the true bandwidth of the signal:

$$\min(2\lambda(\mathcal{T}), f_{\text{nyq}}). \tag{S.3}$$

In most cases, the subband bandwidth $\lambda(\mathcal{T})$ and the actual sampling rate does not exceed f_{nyq} (when not satisfied, the sampling rate is f_{sig}). To ensure reconstruction performance, the parameter p is often not set too low (for instance, it's often chosen to be at least twice the sparsity level $\text{supp}(\mathbf{X})$). It can be seen from above that the theoretical lower bound on the sampling rate is achieved only when $p/LT = pf_s \leq 2\lambda(\mathcal{T}) = 2N_{sig}B$. In other words, when $p = 2\text{supp}(\mathbf{X})$ for the worst case of p, the condition for the actual sampling rate to meet the theoretical lower bound is $|\text{supp}(\mathbf{X})| \leq \frac{N_{\text{sig}}B}{f_s}$. □

II. Proof of Theorem 2

Theorem 2. *When* $r \in [\lceil \frac{f_s}{(M-1)f_s - B}\rceil, \infty)$ *and* $B > f_s$, *we have* $\max_{i\in\{1,\cdots,r\}} \left|supp(\bar{\mathbf{X}}_{S_i})\right| \leq \frac{N_{\text{sig}}B}{f_s}$.

Proof. Consider all blocks $\left\{\mathbf{X}^{U_1}, \cdots, \mathbf{X}^{U_M}\right\}$ in $\mathbf{X}$, where there are consecutive corresponding frequency points of length B. For the case $B > f_s$ and each block occupies M rows in $\mathbf{X}$, the length of the frequency point in each $\bar{\mathbf{X}}^{U_i}$ meets

$$l = B - (M-2)f_s < f_s. \tag{S.4}$$

Because $l < f_s$, we know that the non-zero elements of an any sub-block (i.e. $\bar{\mathbf{X}}^{U_i}$) are distributed on both sides and do not intersect on the columns. Conside one block $\bar{\mathbf{X}}^{U_i}$, We let $r \to \infty$ and observe the change in $\bar{\mathbf{X}}^{U_i}$, $\bar{\mathbf{X}}^{U_i}$ gradually changes from an MMV form to an SMV form. In SMV, the sparsity of the signal is determined by the column in $\bar{\mathbf{X}}^{U_i}$ where it is located. It is observed that

$$\lim_{r\to\infty,i,j} \text{supp}(\bar{\mathbf{X}}^{U_i}_{:,j}) \leq 1. \tag{S.5}$$

Thus, the sparsity of each column in $\bar{\mathbf{X}}^{U_i}$ is less than N_{sig}.

A more complex situation is when r is a finite value, assuming $r = r^{\star}$ is a finite value. In this case, the length of frequency point in each sub-matrix is $\frac{f_s}{r^{\star}}$. We use proof by contradiction to prove the condition that the sparsity of each sub-matrix is less than N_{sig}. Assuming that there exists a sub-matrix $\bar{\mathbf{X}}_{S_i}$ with $\text{supp}(\bar{\mathbf{X}}_{S_i}) > N_{\text{sig}}$. Also, the non-zero elements of an any sub-block of $\bar{\mathbf{X}}_{S_i}$ do not intersect on the columns. $\bar{\mathbf{X}}_{S_i}$ must contain both non-zero elements on both sides of one block. From (S.4), we know that the length of any block of $\bar{\mathbf{X}}$ is less than f_s. We can draw a conclusion that

$$\frac{f_s}{r^{\star}} > f_s - l = (M-1)f_s - B. \tag{S.6}$$

As can be seen, there exist a contradiction between (S.6) and Theorem 2 , so the length of any column-partition sub-matrix non-zero elements in $\bar{\mathbf{X}}$ must be less or equal than $f_s - l$, which is equivalent to r satisfying

$$r \geq \left\lceil \frac{f_s}{(M-1)f_s - B} \right\rceil. \tag{S.7}$$

The proof is thus complete. □

III. Proof of Theorem 3

Theorem 3. *Consider the column-partitioned MMV model (5) with $\min_{i,j}\|(\mathbf{X}_{S_i})_{j,:}\|_2/\|\mathbf{X}_{S_i}\|_F=\eta$ and $|supp(\mathbf{X}_{S_i})|\le s$. Let $s_1:=\min_{i,k}|\Lambda^k_{S_i}\cap supp(\mathbf{X}_{S_i})|$, $s_2:=\min_{i,k}|\Lambda^k_{S_i}\cap supp(\mathbf{X}_{S_i})\cap\tilde{S}^k_{S_i}|$ and $s_3:=\min_{i,k}|\Lambda^k_{S_i}\cap supp(\mathbf{X}_{S_i})\cap\tilde{S}^k_{S_i}\setminus S^k_{S_i}|$. Then, if the sensing matrix $\mathbf{A}$ obeys the RIP with*

$$\delta_{3s}\le\sqrt{\frac{\nu_1\sqrt{\nu_1^2+4\nu_2^2}-\nu_1^2-1}{4\nu_1^2\nu_2^2-2\nu_1^2-1}}\tag{S.8}$$

where $\nu_1:=\frac{1+\omega}{1+\eta\omega\sqrt{s_2}}$ and $\nu_2:=\frac{1+\omega}{1+\eta\omega\sqrt{s_3}}$, SI-SSP produces an signal estimate $\mathbf{X}^k=[\mathbf{X}^k_{S_1},\cdots,\mathbf{X}^k_{S_r}]$ satisfying

$$\left\|\mathbf{X}-\mathbf{X}^k\right\|_F\le\rho^k\left\|\mathbf{X}\right\|_F+\tau\left\|\mathbf{E}\right\|_F,\tag{S.9}$$

where $\rho\in(0,1)$ and τ are constants depending on δ_{3s}, ν_1 and ν_2. Furthermore, after at most $k^*=\lceil\log_\rho\frac{\|\mathbf{X}\|_F}{\tau\|\mathbf{E}\|_F}\rceil$ iterations, SI-SSP estimates $\mathbf{X}$ with

$$\|\mathbf{X}-\mathbf{X}^{k^*}\|_F\le(\tau+1)\|\mathbf{E}\|_F.\tag{S.10}$$

To prove Theorem 3, we first introduce six useful Lemmas, whose proofs are left to the appendices.

Lemma 1. *([25]): For nonnegative numbers a, b, c, d, x, y,*

$$(ax+by)^2+(cx+dy)^2\le(\sqrt{a^2+c^2}x+(b+d)y)^2.\tag{S.11}$$

Lemma 2. *Consider the system model $\mathbf{Y}=\mathbf{AX}+\mathbf{E}$, where $supp(\mathbf{X})=T$ and $|T|=s$. Let $S\subseteq\{1,2,...,n\}$ be an index set with $|S|=t$ and $\mathbf{W}_{T_0}$ be a side-information matrix with diagonal entries indexed by $T_0\subseteq\{1,2,...,n\}$ being $\omega\ge 0$ and zero otherwise. Also, let $\tilde{\mathbf{X}}:=\arg\min_{\mathbf{Z}:supp(\mathbf{Z})\subseteq S}\|\mathbf{Y}-\mathbf{AZ}\|_2$. If $\delta_{3s}<1$, then*

$$\|\mathbf{W}_{T_0}(\mathbf{X}-\tilde{\mathbf{X}})_S\|_F\le\omega\delta_{s+t}\|\mathbf{X}-\tilde{\mathbf{X}}\|_F+\omega\sqrt{1+\delta_t}\|\mathbf{E}\|_F\tag{S.12}$$

and

$$\|\mathbf{X}-\tilde{\mathbf{X}}\|_F\le\sqrt{\frac{1}{1-\delta^2_{s+t}}}\|\mathbf{X}_{S^c}\|_F+\frac{\sqrt{1+\delta_t}}{1-\delta_{s+t}}\|\mathbf{E}\|_F.\tag{S.13}$$

Furthermore, if $t>s$, define T_∇ as the row-indices of the smallest $t-s$ row-norm entries of $\tilde{\mathbf{X}}$ in S, we have

$$\|\mathbf{X}_{T_\nabla}\|_F\le\sqrt{2}\nu_2\delta_{s+t}\|\mathbf{X}-\tilde{\mathbf{X}}\|_F+\nu_2\sqrt{2\left(1+\delta_t\right)}\|\mathbf{E}\|_F.\tag{S.14}$$

Remark 1. *When we consider the atom selection strategy of $\|\tilde{\mathbf{X}}_{T_\nabla}+\mathbf{W}_{T_0}\tilde{\mathbf{X}}_{T_\nabla}\|_F\le\|\tilde{\mathbf{X}}_{S'}+\mathbf{W}_{T_0}\tilde{\mathbf{X}}_{S'}\|_F$, we can also obtain another upper bound for $\|\mathbf{X}_{T_\nabla}\|_F$ in (S.14). In this case, we should allocate $2\left\|\mathbf{X}_{T_\nabla}\right\|_F$ to the left hand side of (A.38), we have*

$$\|\mathbf{X}_{T_\nabla}\|_F\le\sqrt{2}\nu_3\delta_{s+t}\|\mathbf{X}-\tilde{\mathbf{X}}\|_F+\nu_4\sqrt{2\left(1+\delta_t\right)}\left\|\mathbf{E}\right\|_F.\tag{S.15}$$

where $\nu_3=(1-\omega+\omega\delta_{s+t}+\delta_{s+t})/(2\delta_{s+t})$ and $\nu_4=(1+\omega)/(2\delta_{s+t})$.

Lemma 3. *In steps 4 and 5 of SI-SSP, we have*

$$\|\mathbf{X}_{(\tilde{S}^k)^c}\|_F\le\sqrt{2}\nu_1\delta_{3s}\left\|\mathbf{X}-\mathbf{X}^{k-1}\right\|_F+\nu_1\sqrt{2(1+\delta_{3s})}\left\|\mathbf{E}\right\|_F.\tag{S.16}$$

Remark 2. *When we consider the atom selection strategy in select step that*

$$\begin{aligned}&\left\|((\mathbf{I}_L+\mathbf{W}_{T_0})\mathbf{A}^H(\mathbf{Y}-\mathbf{AX}^{k-1}))_T\right\|_F\\&\quad\le\left\|((\mathbf{I}_L+\mathbf{W}_{T_0})\mathbf{A}^H(\mathbf{Y}-\mathbf{AX}^{k-1}))_{\Delta S}\right\|_F.\end{aligned}\tag{S.17}$$

We can also obtain another upper bound for $\|\mathbf{X}_{(\tilde{S}^k)^c}\|_F$ in (S.16). In this case, we should allocate $2\|\mathbf{X}_{(\tilde{S}^k)^c}\|_F$ to the left hand side of (A.51), we have

$$\begin{aligned}\|\mathbf{X}_{(\tilde{S}^k)^c}\|_F\le&\sqrt{2}\nu_4\delta_{3s}\left\|\mathbf{X}-\mathbf{X}^{k-1}\right\|_F\\&+\nu_4\sqrt{2\left(1+\delta_{3s}\right)}\left\|\mathbf{E}\right\|_F.\end{aligned}\tag{S.18}$$

where $\nu_4=(1-\omega+\omega\delta_{3s}+\delta_{3s})/(2\delta_{3s})$ and $\nu_4=(1+\omega)/(2\delta_{3s})$. Based on conclusions (S.15) and (S.18), we know that the sensing matrix $\mathbf{A}$ obeys the RIP with

$$\delta_{3s}\le\sqrt{\frac{\nu_3\sqrt{\nu_3^2+4\nu_4^2}-\nu_3^2-1}{4\nu_3^2\nu_4^2-2\nu_3^2-1}}.\tag{S.19}$$

Lemma 4. *Let $T_0\subseteq\{1,2,...,n\}$, for two vectors $\mathbf{u},\mathbf{v}\in\mathbb{R}^n$, if $|supp(\mathbf{u})\cup supp(\mathbf{v})|\le t$,*

$$\left|\left\langle\mathbf{u},(\mathbf{W}_{T_0}-\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{A})\mathbf{v}\right\rangle\right|\le\omega\delta_t\left\|\mathbf{u}\right\|\left\|\mathbf{v}\right\|;\tag{S.20}$$

Moreover, if $U\subseteq\{1,2,...,n\}$ and $|U\cup supp(\mathbf{v})|\le t$, then

$$\left|(\mathbf{W}_{T_0}-\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{A})\mathbf{v}\right|\le\omega\delta_t\left\|\mathbf{v}\right\|.\tag{S.21}$$

Lemma 5. *For SMV model $\mathbf{y}=\mathbf{\Phi x}+\mathbf{e}$, let $T_0\subseteq\{1,2,...,n\}$, let $U\subseteq\{1,2,...,n\}$ and $|U\cap T_0|\le u$, we have*

$$\left\|(\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{e})_U\right\|_2\le\omega\delta_u\left\|\mathbf{e}\right\|_2.\tag{S.22}$$

Lemma 6. *Consider the MMV model $\mathbf{Y}=\mathbf{AX}+\mathbf{E}$, let $\tilde{\mathbf{X}}$ be the solution of the least squares problem $\arg\min_{\mathbf{Z}}\{\left\|\mathbf{Y}-\mathbf{AZ}\right\|_F, supp(\mathbf{Z})\subseteq S\}$, then*

$$\langle\mathbf{W}_{T_0}\mathbf{X}-\mathbf{W}_{T_0}\tilde{\mathbf{X}},\mathbf{A}^H\mathbf{AZ}\rangle+\omega\left\langle\mathbf{E},\mathbf{AZ}\right\rangle=\mathbf{0}.\tag{S.23}$$

Now we have all ingredients to prove Theorem 3.

Proof of Theorem 3. First, in Steps 4 and 5 of SI-SSP, Lemma 3 implies

$$\|\mathbf{X}_{(\tilde{S}^k)^c}\|_F\le\sqrt{2}\nu_1\delta_{3s}\left\|\mathbf{X}-\mathbf{X}^{k-1}\right\|_F+\nu_1\sqrt{2(1+\delta_{3s})}\left\|\mathbf{E}\right\|_F.\tag{S.24}$$

Note that Step 6 of SI-SSP solves a least squares problem. Let $S=\tilde{S}^k$ and $\tilde{\mathbf{X}}=\tilde{\mathbf{X}}^k$, $t=2s$, by (S.13) we have

$$\|\mathbf{X}-\tilde{\mathbf{X}}\|_F\le\sqrt{\frac{1}{1-\delta^2_{3s}}}\|\mathbf{X}_{(\tilde{S}^k)^c}\|_F+\frac{\sqrt{1+\delta_{2s}}}{1-\delta_{3s}}\|\mathbf{E}\|_F.\tag{S.25}$$

Combining (S.24) and (S.25) and also magnifying δ_{2s} to δ_{3s}, we further have

$$\|\mathbf{X}-\tilde{\mathbf{X}}^k\|_F\le\nu_1\sqrt{\frac{2\delta^2_{3s}}{1-\delta^2_{3s}}}\|\mathbf{X}-\mathbf{X}^{k-1}\|_F+\tau_1\|\mathbf{E}\|_F.\tag{S.26}$$

Next, after Step 7 of SI-SSP, let $S_\nabla=\tilde{S}^k\setminus S^k$ be the row-indices of the smallest $t-s$ row norm entries in $\tilde{\mathbf{X}}^k$. Also, let $T=\tilde{S}^k$, $\tilde{\mathbf{X}}=\tilde{\mathbf{X}}^k$, $T_\nabla=S_\nabla$ and $t=2s$. Then, by (A.37) we have

$$\|\mathbf{X}_{S_\nabla}\|_F\le\sqrt{2}\nu_2\delta_{3s}\|\mathbf{X}-\tilde{\mathbf{X}}^k\|_F+\nu_2\sqrt{2\left(1+\delta_{2s}\right)}\|\mathbf{E}\|_F.\tag{S.27}$$

Let $\tau_1 = (\nu_1\sqrt{2(1-\delta_{3s})}+\sqrt{1+\delta_{3s}})(1-\delta_{3s})^{-1}$ and $\tau_2 = \sqrt{1+\delta_{3s}}$. Dividing $(S^k)^c$ into two disjoint subsets: $(\tilde{S}^k)^c$ and S_∇, we get

$$\begin{aligned}
&\left\|\mathbf{X}_{(S^k)^c}\right\|_F^2 = \|\mathbf{X}_{S_\nabla}\|_F^2 + \|\mathbf{X}_{(\tilde{S}^k)^c}\|_F^2 \\
&\overset{(S.24),(S.27)}{\le} 2\left(\nu_2\delta_{3s}\|\mathbf{X}-\tilde{\mathbf{X}}^k\|_F + \nu_2\tau_2\|\mathbf{E}\|_F\right)^2 \\
&\quad + 2\left(\nu_1\delta_{3s}\left\|\mathbf{X}-\mathbf{X}^{k-1}\right\|_2 + \nu_1\tau_2\left\|\mathbf{E}\right\|_F\right)^2 \\
&\overset{(S.26)}{\le} 2\left(\sqrt{\frac{2\nu_1^2\nu_2^2\delta_{3s}^4}{1-\delta_{3s}^2}}\left\|\mathbf{X}-\mathbf{X}^{k-1}\right\|_F + \nu_2(\tau_1\delta_{3s}+\tau_2)\right. \\
&\quad \left.\times\|\mathbf{E}\|_F\right)^2 + 2\left(\nu_1\delta_{3s}\left\|\mathbf{X}-\mathbf{X}^{k-1}\right\|_F + \nu_1\tau_2\left\|\mathbf{E}\right\|_F\right)^2 \\
&\overset{(S.11)}{\le} 2\left(\sqrt{\frac{2\nu_1^2\nu_2^2\delta_{3s}^4}{1-\delta_{3s}^2}+\nu_1^2\delta_{3s}^2}\left\|\mathbf{X}-\mathbf{X}^{k-1}\right\|_F\right. \\
&\quad \left. + \left((\nu_1+\nu_2)\tau_2+\nu_2\delta_{3s}\tau_1\right)\|\mathbf{E}\|_F\right)^2. \qquad \text{(S.28)}
\end{aligned}$$

Squaring both sides, we get

$$\begin{aligned}
\left\|\mathbf{X}_{(S^k)^c}\right\|_F \le& \sqrt{\tfrac{4\nu_1^2\nu_2^2\delta_{3s}^4}{1-\delta_{3s}^2}+2\nu_1^2\delta_{3s}^2}\left\|\mathbf{X}-\mathbf{X}^{k-1}\right\|_F \\
&+\sqrt{2}\left((\nu_1+\nu_2)\tau_2+\nu_2\delta_{3s}\tau_1\right)\|\mathbf{E}\|_F. \quad \text{(S.29)}
\end{aligned}$$

Step 9 of SI-SSP also solves a least squares problem. Letting $T=S^k$, $\tilde{\mathbf{X}}=\mathbf{X}^k$ and $t=s$, by (S.13), we have

$$\|\mathbf{X}-\mathbf{X}^k\|_F \le \sqrt{\frac{1}{1-\delta_{2s}^2}}\|\mathbf{X}_{(S^k)^c}\|_F + \frac{\sqrt{1+\delta_s}}{1-\delta_{2s}}\|\mathbf{E}\|_F. \tag{S.30}$$

Finally, combining (S.29) and (S.30) yields

$$\|\mathbf{X}-\mathbf{X}^k\|_F \le \rho\|\mathbf{X}-\mathbf{X}^{k-1}\|_F + (1-\rho)\tau\|\mathbf{E}\|_F \tag{S.31}$$

where $\rho := \sqrt{2}\delta_{3s}\sqrt{2\nu_1^2\nu_2^2\delta_{3s}^2+\nu_1^2-\nu_1^2\delta_{3s}^2}(1-\delta_{3s}^2)^{-1}$ and $\tau := \sqrt{2}\delta_{3s}\nu_2(\nu_1\sqrt{2(1-\delta_{3s})}+\sqrt{1+\delta_{3s}})(1-\delta_{3s}^2)^{-1/2}(1-\delta_{3s})^{-1}(1-\rho)^{-1}+(\nu_1\nu_2\sqrt{2(1-\delta_{3s})}+\sqrt{1+\delta_{3s}})(1-\delta_{3s})^{-1}$.

We recursively apply (S.31) to obtain

$$\|\mathbf{X}-\mathbf{X}^k\|_F \le \rho^k\|\mathbf{X}\|_F + \tau\|\mathbf{E}\|_F \tag{S.32}$$

where $\rho<1$ under (S.8). When $k^* = \lceil \log_\rho \frac{\|\mathbf{X}\|_F}{\tau\|\mathbf{E}\|_F}\rceil$, we have $\rho^k\|\mathbf{X}\|_F \le \tau\|\mathbf{E}\|_F$, and thus the stability result (S.10). $\square$

Appendix A
Proof of Lemma 2

- First, we give a upper bound of $\|\mathbf{X}_{T_\nabla}\|_F$, by Lemma 6, let $\mathbf{Z} = (\mathbf{W}_{T_0}\mathbf{X}-\mathbf{W}_{T_0}\tilde{\mathbf{X}})_S$, we have

$$\begin{aligned}
&\left\langle \mathbf{W}_{T_0}(\mathbf{X}-\tilde{\mathbf{X}}), \mathbf{A}^H\mathbf{A}(\mathbf{W}_{T_0}\mathbf{X}-\mathbf{W}_{T_0}\tilde{\mathbf{X}})_S\right\rangle \\
&\quad + \left\langle \mathbf{W}_{T_0}\mathbf{E}, \mathbf{A}(\mathbf{W}_{T_0}\mathbf{X}-\mathbf{W}_{T_0}\tilde{\mathbf{X}})_S\right\rangle = \mathbf{0}. \quad \text{(A.33)}
\end{aligned}$$

Noticing that $\mathrm{supp}(\tilde{\mathbf{X}})\subseteq S$, we have

$$\begin{aligned}
&\|(\mathbf{W}_{T_0}\mathbf{X}-\mathbf{W}_{T_0}\tilde{\mathbf{X}})_S\|_F^2 \\
=\ & \left\langle \mathbf{W}_{T_0}(\mathbf{X}-\tilde{\mathbf{X}}), (\mathbf{W}_{T_0}\mathbf{X}-\mathbf{W}_{T_0}\tilde{\mathbf{X}})_S\right\rangle \\
\overset{(A.33)}{=}\ & \left\langle \mathbf{W}_{T_0}(\mathbf{X}-\tilde{\mathbf{X}}), (\mathbf{I}_L-\mathbf{A}^H\mathbf{A})(\mathbf{X}-\tilde{\mathbf{X}})_S\right\rangle \\
-\ & \left\langle \mathbf{W}_{T_0}\mathbf{E}, \mathbf{A}(\mathbf{W}_{T_0}\mathbf{X}-\mathbf{W}_{T_0}\tilde{\mathbf{X}})_S\right\rangle \\
\overset{(7)}{\le}\ & \omega\delta_{s+t}\left\|(\mathbf{W}_{T_0}\mathbf{X}-\mathbf{W}_{T_0}\tilde{\mathbf{X}})_S\right\|_F\left\|\mathbf{X}-\tilde{\mathbf{X}}\right\|_F \\
& +\omega\|\mathbf{E}\|_F\sqrt{1+\delta_t}\left\|\mathbf{W}_{T_0}(\mathbf{X}-\tilde{\mathbf{X}})_S\right\|_F. \quad \text{(A.34)}
\end{aligned}$$

Divide both sides by $\|(\mathbf{W}_{T_0}\mathbf{X}-\mathbf{W}_{T_0}\tilde{\mathbf{X}})_S\|_F$ to obtain (S.12).

- Next, by expanding [Lemma 2, 25] to the MMV model, we could get a relationship between $\|\mathbf{X}-\tilde{\mathbf{X}}\|_F$ and $\|\mathbf{X}_{S^c}\|_F$. We have

$$\|\mathbf{X}-\tilde{\mathbf{X}}\|_F \le \sqrt{\frac{1}{1-\delta_{s+t}^2}}\|\mathbf{X}_{S^c}\|_F^2 + \frac{\sqrt{1+\delta_t}}{1-\delta_{s+t}}\|\mathbf{E}\|_F. \tag{A.35}$$

- Then, we established the relationship between $\mathbf{X}_{T_\nabla}$ and $\mathbf{X}-\tilde{\mathbf{X}}$. There exist a subset $S'\subseteq S$ and $S'\cap T=\emptyset$. Since T_∇ is defined by the set of indices of the $t-s$ smallest row entries of $\tilde{\mathbf{X}}$, we can conclude that

$$\begin{aligned}
&\|\tilde{\mathbf{X}}_{T_\nabla}\|_F + \|\mathbf{W}_{T_0}\tilde{\mathbf{X}}_{T_\nabla}\|_F \\
\le\ & \|\tilde{\mathbf{X}}_{S'}\|_F + \|\mathbf{W}_{T_0}\tilde{\mathbf{X}}_{S'}\|_F. \quad \text{(A.36)}
\end{aligned}$$

By eliminating the contribution from $T_\nabla\cap S'$ and noticing that $S'\cap T=\emptyset$, we have

$$\begin{aligned}
\|\tilde{\mathbf{X}}_{T_\nabla\setminus S'}\|_F \ &+\ \|\mathbf{W}_{T_0}\tilde{\mathbf{X}}_{T_\nabla\setminus S'}\|_F \\
&\le\ \|(\tilde{\mathbf{X}}-\mathbf{X})_{S'\setminus T_\nabla}\|_F \\
&+\ \|\mathbf{W}_{T_0}(\tilde{\mathbf{X}}-\mathbf{X})_{S'\setminus T_\nabla}\|_F. \quad \text{(A.37)}
\end{aligned}$$

For the left-hand side of (A.37), we have

$$\begin{aligned}
&\|\tilde{\mathbf{X}}_{T_\nabla\setminus S'}\|_F + \|\mathbf{W}_{T_0}\tilde{\mathbf{X}}_{T_\nabla\setminus S'}\|_F \\
&\quad =\ \|(\tilde{\mathbf{X}}-\mathbf{X}+\mathbf{X})_{T_\nabla\setminus S'}\|_F \\
&\quad +\ \|(\mathbf{W}_{T_0}(\tilde{\mathbf{X}}-\mathbf{X})+\mathbf{W}_{T_0}\mathbf{X})_{T_\nabla\setminus S'}\|_F \\
&\quad \ge\ \|\mathbf{X}_{T_\nabla}\|_F + \|\mathbf{W}_{T_0}\mathbf{X}_{T_\nabla}\|_F \quad \text{(A.38)} \\
&\quad -\ \|(\tilde{\mathbf{X}}-\mathbf{X})_{T_\nabla\setminus S'}\|_F \\
&\quad -\ \|\mathbf{W}_{T_0}(\tilde{\mathbf{X}}-\mathbf{X})_{T_\nabla\setminus S'}\|_F. \quad \text{(A.39)}
\end{aligned}$$

Finally, combining (A.39) and (A.37), and noticing that

$$\begin{aligned}
(T_\nabla\setminus S')\cap(S'\setminus T_\nabla) &= \emptyset \quad \text{(A.40)}\\
(T_\nabla\setminus S')\cup(S'\setminus T_\nabla) &\subseteq T, \quad \text{(A.41)}
\end{aligned}$$

we have

$$\begin{aligned}
&\|\mathbf{X}_{T_\nabla}\|_F + \|\mathbf{W}_{T_0}\mathbf{X}_{T_\nabla}\|_F \\
&\quad\le\ \|(\tilde{\mathbf{X}}-\mathbf{X})_{T_\nabla\setminus S'}\|_F + \|(\mathbf{W}_{T_0}(\tilde{\mathbf{X}}-\mathbf{X}))_{T_\nabla\setminus S'}\|_F \\
&\quad+\ \|(\tilde{\mathbf{X}}-\mathbf{X})_{S'\setminus T_\nabla}\|_F + \|(\mathbf{W}_{T_0}(\tilde{\mathbf{X}}-\mathbf{X}))_{S'\setminus T_\nabla}\|_F \\
&\quad\le\ \sqrt{2}\|(\tilde{\mathbf{X}}-\mathbf{X})_{(S'\setminus T_\nabla)\cup(T_\nabla\setminus S')}\|_F \\
&\quad+\ \sqrt{2}\|(\mathbf{W}_{T_0}(\tilde{\mathbf{X}}-\mathbf{X}))_{(S'\setminus T_\nabla)\cup(T_\nabla\setminus S')}\|_F \\
&\quad\le\ \sqrt{2}\|(\tilde{\mathbf{X}}-\mathbf{X})_S\|_F + \sqrt{2}\|(\mathbf{W}_{T_0}(\tilde{\mathbf{X}}-\mathbf{X}))_S\|_F \\
&\quad\overset{(S.12)}{\le}\ \sqrt{2}(1+\omega)\delta_{s+t}\|\mathbf{X}-\tilde{\mathbf{X}}\|_F \\
&\quad+\ (1+\omega)\sqrt{2(1+\delta_t)}\|\mathbf{E}\|_F. \quad \text{(A.42)}
\end{aligned}$$

Also, we can obtain the relationship between $\|\mathbf{W}_{T_0}\mathbf{X}_{T_\nabla}\|_F$ and $\|\mathbf{X}_{T_\nabla}\|_F$:

$$\eta\omega\sqrt{s_3}\,\|\mathbf{X}_{T_\nabla}\|_F \le \|\mathbf{W}_{T_0}\mathbf{X}_{T_\nabla}\|_F. \tag{A.43}$$

Combining (A.42) and (A.43), we have

$$\begin{aligned}\|\mathbf{X}_{T_\nabla}\|_F &\leq \frac{\sqrt{2}(1+\omega)\delta_{s+t}}{1+\eta\omega\sqrt{s_3}}\|\mathbf{X}-\tilde{\mathbf{X}}\|_F \\ &+ \frac{(1+\omega)\sqrt{2(1+\delta_t)}}{1+\eta\omega\sqrt{s_3}}\|\mathbf{E}\|_F. \quad (A.44)\end{aligned}$$

Noting the definition of ν_2, we complete the proof of Lemma 2.

Appendix B
Proof of Lemma 3

Proof: From Step 5 of SI-SSP, we have

$$\mathbf{X}_{S_i}^k = \underset{\mathbf{\Theta}:\mathrm{supp}(\mathbf{\Theta})=S_{S_i}^k}{\arg\min} \|\mathbf{Y}_{S_i}-\mathbf{A}\mathbf{\Theta}\|_F. \quad (A.45)$$

From Step 4 of SI-SSP, let $\mathbf{X}^k=[\mathbf{X}_{S_1}^k,\cdots,\mathbf{X}_{S_r}^k]$. We have the following conclusion

$$\begin{aligned}&\|(\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_T\|_F \\ &+ \|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_T\|_F \\ &\leq \|(\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{\Delta S}\|_F \\ &+ \|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{\Delta S}\|_F. \quad (A.46)\end{aligned}$$

By removing the same coordinates $T\cap\Delta S$, we get

$$\begin{aligned}&\|(\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{T\backslash\Delta S}\|_F \\ &+ \|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{T\backslash\Delta S}\|_F \\ &\leq \|(\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{\Delta S\backslash T}\|_F \\ &+ \|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{\Delta S\backslash T}\|_F. \quad (A.47)\end{aligned}$$

Because $\mathrm{supp}(\mathbf{X})=T$ and $\mathrm{supp}(\mathbf{X}^{k-1})=S^{k-1}$,

$$(\mathbf{X}-\mathbf{X}^{k-1})_{\Delta S\backslash(T\cup S^{k-1})}=0. \quad (A.48)$$

For the right-hand side of (A.47), we have

$$\begin{aligned}&\|(\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{\Delta S\backslash T}\|_F \\ &+ \|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{\Delta S\backslash T}\|_F \\ &= \|(\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{\Delta S\backslash(T\cup S^{k-1})}\|_F \\ &+ \|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{\Delta S\backslash(T\cup S^{k-1})}\|_F \\ &= \|(\mathbf{A}^H(\mathbf{A}\mathbf{X}+\mathbf{E}-\mathbf{A}\mathbf{X}^{k-1}))_{\Delta S\backslash(T\cup S^{k-1})}\|_F \\ &+ \|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{A}\mathbf{X}+\mathbf{E}-\mathbf{A}\mathbf{X}^{k-1}))_{\Delta S\backslash(T\cup S^{k-1})}\|_F \\ &\leq \|((\mathbf{A}^H\mathbf{A}-\mathbf{I}_L)(\mathbf{X}-\mathbf{X}^{k-1}))_{\Delta S\backslash T}\|_F \\ &+ \|(\mathbf{A}^H\mathbf{E})_{\Delta S\backslash T}\|_F \\ &+ \|(\mathbf{W}_{T_0}(\mathbf{A}^H\mathbf{A}-\mathbf{I}_L)(\mathbf{X}-\mathbf{X}^{k-1}))_{\Delta S\backslash T}\|_F \\ &+ \|(\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{E})_{\Delta S\backslash T}\|_F. \quad (A.49)\end{aligned}$$

Note that $\tilde{S}^k=S^{k-1}\cup\Delta S$, we have

$$(\mathbf{X}-\mathbf{X}^{k-1})_{T\backslash\tilde{S}^k}=\mathbf{X}_{(\tilde{S}^k)^c}. \quad (A.50)$$

For the left-hand side of (A.47), we have

$$\begin{aligned}&\|(\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{T\backslash\Delta S}\|_F \\ &+ \|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{T\backslash\Delta S}\|_F \\ &= \|(\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{T\backslash(\Delta S\cup S^{k-1})}\|_F \\ &+ \|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{Y}-\mathbf{A}\mathbf{X}^{k-1}))_{T\backslash(\Delta S\cup S^{k-1})}\|_F \\ &= \|(\mathbf{A}^H(\mathbf{A}\mathbf{X}+\mathbf{E}-\mathbf{A}\mathbf{X}^{k-1}))_{T\backslash(\Delta S\cup S^{k-1})}\|_F \\ &+ \|(\mathbf{W}_{T_0}\mathbf{A}^H(\mathbf{A}\mathbf{X}+\mathbf{E}-\mathbf{A}\mathbf{X}^{k-1}))_{T\backslash(\Delta S\cup S^{k-1})}\|_F \\ &= \|((\mathbf{A}^H\mathbf{A}-\mathbf{I}_L)(\mathbf{X}-\mathbf{X}^{k-1}) \\ &+ \mathbf{A}^H\mathbf{E}+\mathbf{X})_{\Delta S\backslash(T\cup S^{k-1})}\| \\ &+ \|\mathbf{W}_{T_0}(\mathbf{A}^H\mathbf{A}-\mathbf{I}_L)(\mathbf{X}-\mathbf{X}^{k-1}) \\ &+ \mathbf{W}_{T_0}\mathbf{A}^H\mathbf{E}+\mathbf{W}_{T_0}\mathbf{X})_{\Delta S\backslash(T\cup S^{k-1})}\|_F \\ &\geq \|\mathbf{X}_{(\tilde{S}^k)^c}\|_F+\|(\mathbf{W}_{T_0}\mathbf{X})_{(\tilde{S}^k)^c}\|_F \\ &- \|((\mathbf{A}^H\mathbf{A}-\mathbf{I}_L)(\mathbf{X}-\mathbf{X}^{k-1}))_{(\tilde{S}^k)^c}\|_F \quad (A.51) \\ &- \|(\mathbf{W}_{T_0}(\mathbf{A}^H\mathbf{A}-\mathbf{I}_L)(\mathbf{X}-\mathbf{X}^{k-1}))_{(\tilde{S}^k)^c}\|_F \\ &- \|(\mathbf{A}^H\mathbf{E})_{(\tilde{S}^k)^c}\|_F-\|(\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{E})_{(\tilde{S}^k)^c}\|_F. \quad (A.52)\end{aligned}$$

Combining (A.53) and (A.52), we have

$$\begin{aligned}&\|\mathbf{X}_{(\tilde{S}^k)^c}\|_F+\|(\mathbf{W}_{T_0}\mathbf{X})_{(\tilde{S}^k)^c}\|_F \\ &\leq \|((\mathbf{A}^H\mathbf{A}-\mathbf{I}_L)(\mathbf{X}-\mathbf{X}^{k-1}))_{T\backslash\tilde{S}^k}\|_F \\ &+ \|(\mathbf{W}_{T_0}(\mathbf{A}^H\mathbf{A}-\mathbf{I}_L)(\mathbf{X}-\mathbf{X}^{k-1}))_{T\backslash\tilde{S}^k}\|_F \\ &+ \|(\mathbf{W}_{T_0}(\mathbf{A}^H\mathbf{A}-\mathbf{I}_L)(\mathbf{X}-\mathbf{X}^{k-1}))_{\Delta S\backslash T}\|_F \\ &+ \|((\mathbf{A}^H\mathbf{A}-\mathbf{I}_L)(\mathbf{X}-\mathbf{X}^{k-1}))_{\Delta S\backslash T}\|_F \\ &+ \|(\mathbf{A}^H\mathbf{E})_{T\backslash\tilde{S}^k}\|_F+\|(\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{E})_{T\backslash\tilde{S}^k}\|_F \\ &+ \|(\mathbf{A}^H\mathbf{E})_{\Delta S\backslash T}\|_F+\|(\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{E})_{\Delta S\backslash T}\|_F \\ &\leq \sqrt{2}\|((\mathbf{A}^H\mathbf{A}-\mathbf{I}_L)(\mathbf{X}-\mathbf{X}^{k-1}))_{T\cup\Delta S}\|_F \\ &+ \sqrt{2}\|(\mathbf{W}_{T_0}(\mathbf{A}^H\mathbf{A}-\mathbf{I}_L)(\mathbf{X}-\mathbf{X}^{k-1}))_{T\cup\Delta S}\|_F \\ &+ \sqrt{2}\|(\mathbf{A}^H\mathbf{E})_{T\cup\Delta S}\|_F+\sqrt{2}\|(\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{E})_{T\cup\Delta S}\|_F \\ &\leq \sqrt{2}(1+\omega)\delta_{3s}\|\mathbf{X}-\mathbf{X}^{k-1}\|_F \\ &+ (1+\omega)\sqrt{2(1+\delta_{3s})}\|\mathbf{E}\|_F. \quad (A.53)\end{aligned}$$

We can obtain the relationship between $\|\mathbf{X}_{(\tilde{S}^k)^c}\|_F$ $\|(\mathbf{W}_{T_0}\mathbf{X})_{(\tilde{S}^k)^c}\|_F$:

$$\eta\omega\sqrt{s_2}\|\mathbf{X}_{(\tilde{S}^k)^c}\|_F\leq\|(\mathbf{W}_{T_0}\mathbf{X})_{(\tilde{S}^k)^c}\|_F. \quad (A.54)$$

Combining (A.53) and (A.54), we have

$$\begin{aligned}\|\mathbf{X}_{(\tilde{S}^k)^c}\|_F &\leq \frac{\sqrt{2}(1+\omega)\delta_{3s}}{1+\eta\omega\sqrt{s_2}}\|\mathbf{X}-\tilde{\mathbf{X}}\|_F \\ &+\frac{(1+\omega)\sqrt{2(1+\delta_{2s})}}{1+\eta\omega\sqrt{s_2}}\|\mathbf{E}\|_F. \quad (A.55)\end{aligned}$$

Noting the definition of ν_1, we complete the proof of Lemma 3.

Appendix C
Proof of Lemma 4

Proof: the RIC δ_t can be expressed as [25]

$$\delta_t=\max\nolimits_{S\subseteq\{1,2,\ldots,N\},|S|\leq t}\|\mathbf{A}_S^*\mathbf{A}_S-\mathbf{I}\|_{2\to 2}, \quad (A.56)$$

where

$$\|\mathbf{A}_S^*\mathbf{A}_S-\mathbf{I}\|_{2\to 2}=\sup\nolimits_{\mathbf{a}\in\mathbb{R}^{|S|}\backslash\{\mathbf{0}\}}\frac{\|(\mathbf{A}_S^*\mathbf{A}_S-\mathbf{I})\,\mathbf{a}\|_2}{\|\mathbf{a}\|_2}. \quad (A.57)$$

Let $S=\mathrm{supp}(\mathbf{u})\cup\mathrm{supp}(\mathbf{v})$, then $|S|\leq t$. Let $\mathbf{u}_{|S},\mathbf{v}_{|S}$ denote respectively the S-dimensional sub-vectors of $\mathbf{u}$ and

$\mathbf{v}$ obtained by only keeping the components indexed by S. We have

$$\begin{aligned}
& |\langle \mathbf{u}, \left(\mathbf{W}_{T_0} - \mathbf{W}_{T_0}\mathbf{A}^H\mathbf{A}\right)\mathbf{v}\rangle| \\
= \quad & |\langle \mathbf{W}_{T_0}\mathbf{u}, \mathbf{v}\rangle - \langle \mathbf{A}\mathbf{W}_{T_0}\mathbf{u}, \mathbf{A}\mathbf{v}\rangle| \\
= \quad & |\langle \mathbf{W}_{T_0}\mathbf{u}_{|S}, \left(\mathbf{I}_L - \mathbf{A}_S^H\mathbf{A}_S\right)\mathbf{v}_{|S}\rangle| \\
\leq \quad & \|\mathbf{W}_{T_0}\mathbf{u}_{|S}\|_2 \left\|\left(\mathbf{I}_L - \mathbf{A}_S^H\mathbf{A}_S\right)\mathbf{v}_{|S}\right\|_2 \\
\overset{(A.57)}{\leq} \quad & \|\mathbf{W}_{T_0}\mathbf{u}_{|S}\|_2 \|\mathbf{I}_L - \mathbf{A}_S^H\mathbf{A}_S\|_{2\to 2}\|\mathbf{v}_{|S}\|_2 \\
\overset{(A.56)}{\leq} \quad & \omega\delta_t\|\mathbf{u}_{|T}\|_2\|\mathbf{v}_{|S}\|_2 \\
= \quad & \omega\delta_t\|\mathbf{u}\|_2\|\mathbf{v}\|_2,
\end{aligned} \tag{A.58}$$

moreover, we have

$$\begin{aligned}
& \|\left(\left(\mathbf{W}_{T_0} - \mathbf{W}_{T_0}\mathbf{A}^H\mathbf{A}\right)\mathbf{v}\right)_U\|_2^2 \\
= \quad & \langle\left(\left(\mathbf{W}_{T_0} - \mathbf{W}_{T_0}\mathbf{A}^H\mathbf{A}\right)\mathbf{v}\right)_U, \\
& \left(\mathbf{W}_{T_0} - \mathbf{W}_{T_0}\mathbf{A}^H\mathbf{A}\right)\mathbf{v}\rangle \\
\overset{(S.20)}{\leq} \quad & \delta_t \left\|\left(\left(\mathbf{W}_{T_0} - \mathbf{A}^H\mathbf{A}\right)\mathbf{v}\right)_U\right\|_2 \|\mathbf{v}\|_2
\end{aligned} \tag{A.59}$$

which completes the proof of Lemma 4.

Appendix D
Proof of Lemma 5

Proof: The lemma follows trivially from the fact that

$$\begin{aligned}
& \left\|\left(\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{e}\right)_U\right\|_2^2 \\
= & \left\langle \mathbf{W}_{T_0}\mathbf{A}^H\mathbf{e}, \left(\mathbf{W}_{T_0}\mathbf{A}^H\mathbf{e}\right)_U\right\rangle \\
= & \left\langle \mathbf{e}, \mathbf{W}_{T_0}\mathbf{A}\left(\left(\mathbf{A}^H\mathbf{e}\right)_U\right)\right\rangle \\
\leq & \|\mathbf{e}\|_2 \left\|\mathbf{W}_{T_0}\mathbf{A}\left(\left(\mathbf{A}^H\mathbf{e}\right)_U\right)\right\|_2 \\
\overset{(7)}{\leq} & \|\mathbf{e}'\|_2\, \omega\sqrt{1+\delta_u}\left\|\left(\mathbf{A}^H\mathbf{e}\right)_U\right\|_2 .
\end{aligned} \tag{A.60}$$

Appendix E
Proof of Lemma 6

Proof: Due to the orthogonality, the residue $\mathbf{Y} - \mathbf{A}\tilde{\mathbf{X}}$ is orthogonal to the space $\mathbf{A}\mathbf{Z}$. This means that for all $\mathbf{Z} \in \mathbb{C}^{L\times N}$ with $\text{supp}(\mathbf{Z}) \subseteq S$,

$$\langle \mathbf{A}(\mathbf{Y} - \mathbf{A}\tilde{\mathbf{X}}), \mathbf{Z}\rangle = \mathbf{0}. \tag{A.61}$$

Let $\tilde{\mathbf{X}}'$ be the solution of the least squares problem $\arg\min_{\mathbf{Z}}\{\|\mathbf{Y}' - \mathbf{A}\mathbf{Z}\|_F, \text{supp}(\mathbf{Z}) \subseteq S\}$, where $\mathbf{Y}' = \frac{\mathbf{A}\mathbf{W}_{T_0}\mathbf{X}_{T_0}}{\omega} + \mathbf{E}$. We have

$$\tilde{\mathbf{X}}' = \frac{\mathbf{W}_{T_0}\tilde{\mathbf{X}}}{\omega}. \tag{A.62}$$

Then, by (A.61), we have

$$\begin{aligned}
\mathbf{0} \quad = \quad & \left\langle \frac{\mathbf{A}\mathbf{W}_{T_0}\mathbf{X}_{T_0}}{\omega} + \mathbf{E} - \mathbf{A}\frac{\mathbf{W}_{T_0}\tilde{X}}{\omega}, \mathbf{A}\mathbf{Z}\right\rangle \\
= \quad & \left\langle \mathbf{W}_{T_0}\mathbf{X} - \mathbf{W}_{T_0}\tilde{\mathbf{X}}, \mathbf{A}^H\mathbf{A}\mathbf{Z}\right\rangle + \omega\left\langle \mathbf{E}, \mathbf{A}\mathbf{Z}\right\rangle .
\end{aligned} \tag{A.63}$$